STEALING THE RICHES: USING THE HUMAN GENOME PROJECT FOR LIVESTOCK RESEARCH

CDK Bottema[1]
[1]Dept Animal Science, Livestock Systems Alliance, Adelaide University, Roseworthy Campus, SA 5371

Abstract
The human genome project has brought a new era not only to medical genetics, but also to livestock molecular genetics. By "borrowing" everything from techniques and research strategies to actual data from the human genome project, geneticists are making significant progress in applying DNA biotechnologies to livestock production. These applications include parentage testing, identity testing, and diagnostic testing of genetic disorders. Another important application is selection of superior animals for breeding programs by identifying those carrying specific genes. Examples of these applications will be discussed from our work on cattle and sheep molecular genetics.

Introduction

The past 10 years have seen the advent of DNA biotechnologies as techniques have been developed and knowledge gained. These technologies include transgenesis, gene therapy, genetic disease carrier testing, and DNA fingerprinting to name but a few. Because of the medical and forensic applications of DNA biotechnologies, the field of human molecular genetics has led the way in these developments. However, scientists in agricultural research have been able to capitalise on these developments to further their own endeavours.

Consequently, the DNA biotechnologies are now being used extensively in the production of crops. These applications include the selection of superior plants for breeding programs, the identification of specific plant cultivars, and the genetic engineering of new plant varieties.

Unfortunately, these DNA biotechnologies are not being applied to the same extent in livestock production. On the other hand, livestock research has been able to more directly exploit the tremendous wealth of results from the human genome project.

Human genome project

So what are the applications of the human genome project for livestock research? The applications fall into 2 categories: 1) applications that utilise the data generated from the project, and 2) applications that utilise the technologies developed for the project. The major application of the data from the human genome project is the identification of genes that control production or disease traits in livestock. The major application of the technology from the human genome project is DNA-based diagnostic testing.

To understand role of the human genome project, perhaps it is best to take a step back and describe the knowledge and techniques necessary for these applications.

Identifying and characterising genes

There are many reasons that livestock scientists are interested in identifying genes involved in the livestock production. These include basic scientific knowledge, genotypic selection and phenotypic modification. The ultimate goal, though, is to be able to manipulate traits.

The problem is that in order to identify the particular gene controlling a specific trait, researchers must search through ~3 x 10^9 DNA bases in the genome. This is a big task because i) the functional coding part of genes represents only 7% of mammalian genomes (2.2 x 10^8 bases out of 3 x 10^9 bases), and ii) there are no true landmarks to distinguish genes from the extragenic sequences (so-called "junk DNA").

Fortunately, as a result of the human genome project, nearly the entire human DNA sequence is available. Using expression data and putative landmarks, the initial search of the human sequence data identified approximately 30,000 genes. However, it is now clear that not all human genes had been located and recent estimates are 80,000 genes for mammalian genomes.

Most importantly, the majority of the human genes located thus far are anonymous (that is, the function of the gene is unknown). Of the 80,000 genes, only about 10,000 genes (<15%) have been thoroughly analysed and characterised. Thus, the search for a particular gene controlling a specific trait is still like "looking for a needle in the proverbial haystack".

Identification of livestock genes

Given the difficulties of identifying genes, the most successful strategy for livestock geneticists has been the positional comparative candidate gene approach. Simply stated, the steps of this strategy are:

1) locate or map the trait to a region of the livestock genome,
2) examine the equivalent region from the human genome for genes that are the most likely candidates to be involved with the trait, and
3) clone and characterise the livestock candidate genes.

Thus, the first step defines the position of the gene in the livestock genome, the second step compares this position in other species to implicate possible genes, and the third step verifies that the gene is

correct. Each step is arduous, costly, and time-consuming. So as may be imagined, only a handful of livestock genes have been associated with traits.

Genome mapping

The first step for the positional comparative candidate gene approach is gene mapping. Gene mapping is the localisation of genes within a genome of a species. Gene mapping implies:

- the chromosome is known,
- the approximate position of the gene on the chromosome is known,
- the distance between the gene and other genetic loci on the chromosome is known, and
- the order of the gene is known with respect to the other genetic loci.

Genome mapping is a more general term for the localisation of genetic loci because it includes more types of loci than just genes. The types of loci localised by genome mapping include genes, markers, and traits.

Markers are usually unique sequence variants that serve as signposts so the researcher can distinguish between the regions of the genome. The markers can be genes, but more often markers are extragenic, anonymous DNA sequences.

There are two types of genome mapping: genetic linkage mapping, and physical mapping. Physical mapping determines the chromosomal location, while genetic linkage mapping determines the genetic distance and order of the loci.

Genetic linkage

Genetic linkage means that two loci (eg. markers or genes) are usually inherited together. Linkage occurs because two loci are close to each other on the same chromosome. Because the two loci are near each other (linked), they are inherited together in the next generation. Genetic linkage mapping is the analysis of loci inheritance within appropriate families.

If a gene controlling a trait is close to a marker, then the trait and marker are linked and will be usually inherited together. Therefore, the marker can be used to follow the inheritance of the trait. If the marker has been mapped, then the genetic loci controlling the trait can be located on the genome map since it must be near the marker. Thus, traits can be mapped as well as genes and markers.

To locate or map the region of the genome controlling a trait, consequently, involves following the inheritance of the trait with markers distributed throughout the genome. The region of the genome that contains a genetic locus controlling a trait is called a quantitative trait locus (QTL). The QTL information can be used for 2 purposes. Firstly, unique markers near the QTL can be used to select for the trait by marker-assisted selection (MAS) (see below). Secondly, the QTL data are fundamental to identifying the actual gene by comparative mapping to other species or by cloning approaches.***Comparative mapping***

Unfortunately, mapping the trait to a region of a livestock genome is not sufficient to directly identify the gene since the region will represent more than one million bases. Since the livestock genome maps have only hundreds of genes located, it is highly unlikely that the identity of gene can be deduced straight from the mapping data.

Therefore, instead of using the "gene-poor" livestock genome maps, researchers rely upon the "gene-rich" genome maps of human and mouse. This is possible because a group of genes that are physically linked together in a human chromosomal region is usually conserved as the same linkage group in other mammalian species.

Thus, having found the region of the livestock genome controlling a trait, that region is aligned to the human and mouse gene maps. By finding the equivalent or homologous region on the human or mouse map, one may determine which genes on the human or mouse map are the most likely candidates to control the trait in livestock.

In general, the human and livestock maps are very similar with large regions in common. The mouse map is more disjoint, making comparisons more difficult. Consequently, candidate genes for livestock traits are typically found by examining the human sequence and mapping databases.

The main consideration is that for any given region of the human genome not all genes will have been identified, let alone characterised. Therefore, any physiological, metabolic or biochemical clues are of great importance. If a good candidate is nominated, then the inheritance of that gene with the trait will be studied in different populations. Should the candidate appear tightly linked or associated with the trait, then the gene can be fully characterised in hopes of identifying the sequence variant responsible for the trait.

Example: spider lamb syndrome

A classic example of this approach was the identification of the gene causing spider lamb syndrome in Suffolk and Hampshire sheep. This lethal recessive genetic defect results in chondrodyplasia and “spider-like” appearance of lambs. The defect has been reported in the United States and Australia. Carrier testing for this defect is essentially as heterozygotes are not distinguishably by phenotype (SS = affected, NN and NS = normal).

Fortunately, the spider lamb gene was mapped in sheep and it was noted that the location of the spider lamb gene was equivalent to a region of the human genome containing a gene (FGFR3) causing a type of chondrodyplasia in humans. The gene was subsequently sequenced in Suffolk sheep and the mutation discovered. The same mutation was observed both in the United States and Australia so a single DNA diagnostic test for carriers could be developed.

Example: cattle gene mapping project

There are many gene mapping projects for livestock species worldwide. At the University of Adelaide, we established a large-scaled gene mapping project in cattle. The 2 major aims of the project are: i) to study the mode of inheritance of important meat quality traits, and ii) to map major genes controlling these traits. Once it is understood how traits are inherited, then improved selection strategies will be available for producer breeding programs. Mapping the major genes will lead to their identification for study and manipulation.

Scientifically, the five overall research aims of the Adelaide Cattle Gene Mapping Project for cattle production traits are:

i) improve the bovine physical gene map for comparisons to the human map,

ii) determine the mode of inheritance of production traits in cattle (specifically meat quality),

iii) map the traits in an experimental herd,

iv) locate and identify those genes involved in controlling the traits of interest, and

v) develop marker-assisted selection strategies for the traits.

The Adelaide Cattle Gene Mapping Project is part of an international effort to map the cattle genome, and complements the other beef cattle gene mapping programs. In fact, during 1995, AgResearch established a sister herd to the Adelaide Gene Mapping herd in New Zealand. Consequently, the Adelaide-AgResearch Cattle Gene Mapping Project now has several important features that distinguish the Project from all other current beef cattle mapping work. These include:

- the breeds (Jersey and Limousin),
- the herd design (double backcross),
- the herd size (one of largest in the world with almost 800 gene mapping progeny),
- the only herd in two environments (Australia and New Zealand),
- the only herd with two feeding regimes (grain and grass finishing), and
- the number of phenotypic traits that can be measured (breeds differ significantly in > 20 traits), and
- the additional measurements (fat and protein metabolic traits).

The traits that are being mapped fall into 11 broad categories:growth traits

- skeletal traits
- age of puberty
- carcass measurements
- yield (full bone-out)
- fat traits
- tenderness
- hide traits
- temperament
- feed intake (Australia only), and
- disease resistance (NZ only).

For each of these categories, QTL have been mapped and, using the data from human genome project, work is in progress to identify the genes themselves. In addition, there is the opportunity to compare the results of grain feeding versus pasture on beef carcass traits.

Livestock DNA biotechnology applications

Given the human genome project and the intense efforts of livestock scientists worldwide, it may appear surprising that DNA biotechnologies are currently under-utilised by the livestock industries. Remarkable advances have been made in human molecular genetics as the result of much hard work by a large number of scientists and substantial financial resources.

Why have there not been the equivalent advances in plant and livestock molecular genetics? The simple answer is that there have been fewer scientists and less financial resources for agricultural based research. Fortunately for those of us in agricultural research, we can adapt the techniques and knowledge from the human work for plant and livestock molecular genetics. So although there is a significant lag in agricultural molecular genetics research, eventually the gap will close if the resources are provided.

As eluded above, sufficient progress has been made to permit the application of some DNA biotechnologies to a few agricultural systems, particularly the production of crops. Why are there fewer applications of DNA biotechnologies in livestock production? Again, it is a matter of financial resources and time. Experimental work in livestock is usually more expensive than the equivalent work in crops because of the costs associated with producing and maintaining large animals. The generation of time of animals is also longer than most crops, so more time is required to complete the same work. The exception, of course, would be crops such as fruit trees where even more time is required.

It is predictable then that those livestock DNA biotechnology applications that are available utilise several fundamental techniques adapted from medical work. Currently, the applications range from relatively simple DNA tests to the technically difficult gene engineering. However, the power of DNA biotechnology in livestock production is only beginning to be realised, and new techniques and applications are always on the horizon.

New technologies

To complete the human genome project, it has been necessary to develop a host of molecular biology techniques which allow researchers to find genes within a genome, characterise the organisation

of these genes, determine their sequence, elucidate their function, and ultimately, manipulate their products. These techniques are highly specialised and expensive. However, these new techniques have not only accelerated the rate of genomic research, but have also enabled the study and manipulation of genes once they have been identified.

Among the fundamental molecular biology techniques are the polymerase chain reaction (PCR), gene cloning, Southern blot hybridisation, sequencing, gene expression systems, and gene transfer. Some of these techniques (eg. PCR, cloning) allows the researcher to produce many copies (amplify) of a specific DNA sequence. Once sufficient material is available, other techniques allow the researcher to study and characterise the DNA (eg. Southern blot hybridisation, sequencing, gene expression systems, gene transfer).

DNA diagnosis and testing

The same molecular biology techniques used to study and characterise DNA can be adapted for biotechnology applications. The technique chosen for any particular application is usually the one that is least expensive but still robust. For instance, once it has been determined that a specific mutation within a gene causes a genetic disorder, it is possible to detect the DNA mutation in affected individuals, affected embryos or carriers by using PCR, Southern blot hybridisation or sequencing. Identical techniques can be used to detect pathogens if a DNA sequence is unique to particular pathogen. In these scenarios, PCR will be usually the method of testing because it the least expensive and technically difficult.

An example would be the test for the spider lamb syndrome (see above). With the discovery of the mutation, it was possible to develop a test to detect carriers of this genetic disorder. Thus, a prized ram carrying the disorder can be still mated provided only those progeny not carrying the disorder are retained.

These techniques can be also applied to other types of DNA testing. For instance, if a specific sequence is unique to males (eg. DNA sequences on the Y chromosome in mammals), then the sex of individuals (eg. certain types of sexually indistinguishable birds) or embryos can be determined. DNA testing can be used to identify particular breeds, strains or species if specific DNA sequences are known for these breeds, strains or species.

In addition, DNA testing can determine the level of genetic diversity in a breed, strain or species. In this case, the variation in DNA sequences is studied between different populations. Obviously, the variation does not usually occur in the genes since this would be likely to interfere with gene function. However, the "junk" DNA sequence is not so constrained, and can vary. The greater the DNA variation, the more genetic diversity is present. This is important in conservation biology because if individuals are genetic diverse (that is, not related), the progeny are more fit.

Marker-assisted selection

Another approach to improve the progeny is to use marker-assisted selection in breeding programs. In marker-assisted selection (MAS), DNA markers are identified which are associated or linked to superior genes. If a DNA marker is linked to a superior gene, then it will be usually inherited with the gene. An example in cattle is the GeneStar Marbling™ marker.

Molecular biology techniques such as PCR or Southern blot hybridisation are used to determine which individuals are carrying which markers. These markers are then used as the basis of selection assuming the superior genes will be also inherited.

By coupling MAS with normal phenotypic selection in a breeding program, one is more likely to obtain the desired results. MAS can be greatly improved by identifying the genes themselves and determining the specific DNA change that gives the superior genotype. When this is possible, one can then perform direct DNA testing for the superior genes.

DNA fingerprinting

Another form of DNA testing is fingerprinting. DNA fingerprinting exploits the variation between individuals within a species. Among the "junk" DNA sequences that do not encode genes are different types of DNA repeats (microsatellites). Although the reasons are not entirely clear, these DNA repeats tend to vary between individuals because the number of repeats at any specific location in the genome mutates occasionally. On the other hand, most of the time, an individual will have the same number of repeats at a given location as their parents because the number of repeats is usually stably inherited.

Since the molecular biology techniques (eg. Southern blot hybridisation) can determine the number of repeats at a given site within the genome, the variation between individuals can be observed. By observing many sites simultaneously, a distinctive pattern (DNA fingerprint) emerges. Thus, individuals can be identified. Moreover, by observing which repeats are inherited at the different sites, parentage can be established.

Gene engineering, transgenesis and gene therapy

Molecular biology techniques can be applied one step further to actually produce a gene product. Cloning a gene allows more DNA to be made. However, by engineering or modifying a gene clone, it possible to have the DNA express or produce a gene product, usually a protein. The protein can be produced in cultured bacteria, yeast or mammalian cells, purified and used for medical applications or other purposes.

If the gene is engineered and transferred to whole plant or animal instead of cultured cells, the

gene is called a transgene and the process is referred to as transgenesis. If the gene is engineered and transferred to a human, the process is referred to as gene therapy. This is because transgenesis and gene therapy serve very different purposes.

Transgenesis is used to produce large quantities of a particular protein or to improve plants and animals. To produce particular proteins, a transgene which encodes the protein is introduced into a plant or animal and the protein is harvested (eg. production of vaccines in sheep milk). To improve plants or animals, a transgene which encodes an improved protein or more protein is introduced (eg. growth hormone in pigs). In transgenesis, it possible for the transgene to be introduced in the germline and inherited by the progeny.

In gene therapy, genes are introduced into humans to treat genetic defects or defects caused by cancer. Thus, gene therapy is used to provide a medical patient who lacks a specific gene with the correct functional gene. Gene therapy is not introduced into the germline of the patient and cannot be inherited by their children.

Current state of DNA biotechnologies in livestock

What is the current state of DNA biotechnology applications in livestock production? The costs and technical challenges associated with many of the DNA biotechnologies prohibit their widespread use. Some of the techniques (eg. PCR) are relatively inexpensive and the costs will continue to decrease as robotic automation and miniaturisation becomes possible. Other techniques are labour intensive (eg. Southern blot hybridisation), while others use expensive reagents (eg. sequencing) or are technically demanding (eg. gene transfer). Thus, simple DNA tests by PCR are more commonly used than transgenesis.

The lack of knowledge about genes in livestock species also hinders widespread use of many DNA biotechnologies. Some of the applications are clearly under-employed. For instance, as mentioned above, even in the well-studied human genome <15% of the genes have been characterised. In most livestock species, less than 1000 genes (2% of the genes) have been analysed. Consequently, DNA diagnosis is limited to a few examples (eg. porcine stress syndrome, cattle double muscling).

On the other hand, other techniques can be readily applied to livestock production. For example, DNA repeats have been sufficiently characterised to permit very accurate DNA fingerprinting in many mammalian species. The technology has improved so that Southern analysis has been replaced with cheaper PCR methods. Correspondingly, commercial services are becoming available for DNA pedigreeing and identity testing for most animals.

Future of DNA biotechnology in livestock

Given the lack of knowledge and the high costs, it is no wonder that DNA biotechnologies are not playing a large role in livestock production. It is likely that some DNA biotechnologies will have more of an impact than others, at least in the short term. What can we realistically expect?

DNA diagnostics

For some areas, such as DNA diagnostics, many tests are currently in use but are limited to the stud breeding industries. Since the cost of DNA diagnosis should decrease from the present costs of approximately $50 per test to $5 per test, the use of DNA diagnostics should continue to increase. Commercial producers as well as stud breeders should be able to afford routine DNA testing.

However, because of the nature of DNA tests in livestock, any given genetic test will have a limited lifespan as the animals carrying the undesirable genes are eliminated. Therefore, the increased use of DNA testing will only occur if new genetic disease mutations are identified or if beneficial DNA changes giving superior genes are identified.

Marker-assisted selection

Marker-assisted selection suffers from similar problems. Few useful markers have been identified in livestock species. Moreover, these markers are usually associated with a simple trait controlled by single gene (eg. presence of horns, coat colour). Most traits (eg. milk yield, meat yield, fleece weight) are controlled by many genes. Thus, until many more markers have been associated with many more traits, marker-assisted selection will be of limited use in livestock.

It will also be necessary for the cost of DNA technologies to decrease if marker-assisted selection is to be commercially viable. As with the DNA diagnosis, this should be possible.

DNA fingerprinting

DNA fingerprinting is presently too expensive ($30-$50 per test) to have found widespread acceptance in the livestock production. Again it is limited to the breeding stud industries despite the usefulness of the results. Many new approaches (eg. single nucleotide polymorphism (SNP) genotyping and DNA chips) are now being investigated to determine if they can be used for parentage determination and individual identification. These new approaches are much less expensive ($5 per test) and are likely to supersede the current DNA fingerprinting methodologies.

We have developed prototype parentage testing panels using single base changes (single nucleotide polymorphisms, SNPs) for cattle and sheep because the technology to detect genotype the SNP genotypes from the human genome project can be minituarised and automated. Thus, genotyping SNPs is significantly less expensive than the current technologies for genotyping repeats.

If these new DNA fingerprinting approaches are successful, they may revolutionise livestock production. Commercial producers will be able to afford to accurately determine parentage within their herds or flocks and can improve their breeding programs. End-products, such as a beef steak, can be traced back from the meat processor to the abattoir to the feedlot to the original producer so the producer can be rewarded or improve production.

Gene engineering

Gene engineering using various cell culture systems is widely used by biotechnology companies commercially. This is likely to continue as long as the products cannot be produced by other means. The products from gene engineering are usually more relevant in terms of medical applications, etc. than in terms of livestock production.

However, gene engineering may have other roles in livestock production. For example, it is possible to introduce engineered genes into ruminant bacteria to improve production of cattle, goats and sheep. The difficulty is that the appropriate genes to engineer must be first identified and methods of transfer developed.

Transgenesis

It is reasonable to assume that the expense and technical difficulties presently plaguing livestock transgenesis will continue to limit its present role. Transgenic animals will continue to be invaluable research tools. Transgenic animals will also continue to play a major role for pharmaceutical companies who can afford to spend millions of dollars to create a single animal capable of producing a specific hormone, vaccine, or anti-snake venom.

In contrast, until new techniques are developed to readily transfer genes into livestock, it is unlikely that genetically modified livestock strains will be available commercially. However, techniques for transferring genes in mice have been reasonably successful, and if these techniques can be improved or adapted for livestock, this can rapidly change. Likewise, if the efficiency of cloning can be improved, then the production of genetically modified livestock may be more feasible.

Conclusions

It is sometimes frustrating for animal scientists to observe the progress in medical genetics and not to be able to offer the same applications of DNA biotechnologies for livestock production. On the other hand, animal scientists are buoyed by the fact that these DNA biotechnologies are being applied successfully in crop production and that there is no reason to assume they cannot be applied successfully to livestock production. It is just a matter of time.